

```

-----
Helix Systems -- Center for Information Technology
http://helix.nih.gov
National Institutes of Health, Department of Health and Human Services

DNAWorks Web Site: http://helixweb.nih.gov/dnaworks
-----

```

```

-----
Send all correspondence to webtools@helix.nih.gov
-----

```

Job started on 06/10/2016 at 16:29:09

Job name: MDQ061016NmeningitidisNacetylNeuraminicAcidSynthase

Output will be sent to myquan929@gmail.com

SEQUENCE 1: PROTEIN LENGTH = 450

```

-----
1 MEKVNEERDAVFEDHIGDRRRSVRSLLLEAFADMEKTSYDVEVADTPQPHIPIRFRHPP
61 IAGPVHDFVGDIAHDFQKMMKRGQAVDFCHWVSHLIATEIDEKFSEVAFRDVQYNPDY
121 VTDSTTEAKKLFNDKIWPAIDKILQQAETCPILSEKWSGIHVSGDQLKGQRHKQEDRFL
181 AYPNGQYMDRGEDPISVLAVFDGHHGHECSQYAAGHLWETWLEVRKSRDPSDSLEDQLRK
241 SLELLDERMTVRSVKECWKGGSTAVCCAIDMDQKLMALAWLGDSPGYVMSNIEFRQLTRG
301 HSPSDEREARRVEEAGQLFVIGGELRVNGVLNLTRALGDVPGRPMISNEPETCQVPIES
361 SDYLVLLACDGISDVFNERDLYQLVEAFANDYPVEDYAELSRFICTKAIEAGSADNVSVV
421 IGFLRPPQDVWKLKMHESDDESDVTDEEX
-----

```

#### CODON FREQUENCY TABLE: E. coli Class II

```

-----
TTT F 0.291   TCT S 0.324   TAT Y 0.352   TGT C 0.389
TTC F 0.709   TCC S 0.266   TAC Y 0.648   TGC C 0.612
TTA L 0.034   TCA S 0.048   TAA X 0.630   TGA X 0.352
TTG L 0.055   TCG S 0.074   TAG X 0.076   TGG W 1.000

CTT L 0.056   CCT P 0.112   CAT H 0.298   CGT R 0.643
CTC L 0.080   CCC P 0.016   CAC H 0.702   CGC R 0.330
CTA L 0.008   CCA P 0.153   CAA Q 0.187   CGA R 0.011
CTG L 0.767   CCG P 0.719   CAG Q 0.814   CGG R 0.008

ATT I 0.335   ACT T 0.291   AAT N 0.173   AGT S 0.045
ATC I 0.659   ACC T 0.536   AAC N 0.828   AGC S 0.243
ATA I 0.006   ACA T 0.047   AAA K 0.786   AGA R 0.006
ATG M 1.000   ACG T 0.127   AAG K 0.215   AGG R 0.003

GTT V 0.398   GCT A 0.275   GAT D 0.461   GGT G 0.508
GTC V 0.135   GCC A 0.161   GAC D 0.540   GGC G 0.428
GTA V 0.200   GCA A 0.240   GAA E 0.754   GGA G 0.020
GTG V 0.268   GCG A 0.323   GAG E 0.247   GGG G 0.044

```

#### ACTIVE CODONS FOR SEQUENCE GENERATION

Residue	Codons	Active Codons	Min. Freq.	Max. Freq.
A Ala	4	4	0.161	0.323
C Cys	2	2	0.389	0.612
D Asp	2	2	0.461	0.540
E Glu	2	2	0.247	0.754
F Phe	2	2	0.291	0.709
G Gly	4	2	0.428	0.508

H	His	2	2	0.298	0.702
I	Ile	3	2	0.335	0.659
K	Lys	2	2	0.215	0.786
L	Leu	6	2	0.080	0.767
M	Met	1	1	1.000	1.000
N	Asn	2	2	0.173	0.828
P	Pro	4	3	0.112	0.719
Q	Gln	2	2	0.187	0.814
R	Arg	6	2	0.330	0.643
S	Ser	6	3	0.243	0.324
T	Thr	4	3	0.127	0.536
V	Val	4	4	0.135	0.398
W	Trp	1	1	1.000	1.000
X	End	3	2	0.352	0.630
Y	Tyr	2	2	0.352	0.648

# SEQUENCE PATTERNS TO BE SCREENED

## PARAMETERS FOR TRIAL 1

Total Size Of Gene ..... 1350 nt  
 Protein Residues ..... 450  
 Mutatable Residues ..... 431  
 Fixed Nucleotides ..... 57 nt  
 Degenerate Nucleotides ..... 0 nt  
 Oligo Size ..... 60 nt  
 Annealing Temp ..... 62 +/- 1°C  
 Oligo Concentration ..... 1.00E-7 M  
 Sodium Concentration ..... 5.00E-2 M  
 Mg2+ Concentration ..... 2.00E-3 M  
 Codon Frequency Threshold .. 10%  
 Repeat Threshold ..... 8 nt  
 Mispriming Threshold ..... 8/18 (6 exact) nt

The DNA sequence # 1 is:

1 ATGGA AAAAGTTAATGAAGAACGTGACGCGGTGTTTGAAGATCACATCGGCGACCGTCGC  
 61 CGCTCTGTTCTGTTCTCTGCTCGAAGAAGCTTTTCGCGGACGAAATGGAGAAAACCTCTTAC  
 121 GACGTTGAGGTTGCGGATACCCACAGCCGCACATCCCAATCCGTTTTCGCCACCCGCCG  
 181 ATCGCTGGCCAGTTCATGATGTGTTTCGGTGATGCTATCCACGACATCTTTCAGAAAATG  
 241 ATGAAACGCGGTGAGGCTGTTGACTTCTGCCACTGGGTGTCTCACCTGATTGCGACTGAA  
 301 ATCGACGAAAAGTTTTCCGAAGTTGCGTTCCGTGACGTTTCACTACAACCCGACATCTAC  
 361 GTTACGGA CTACTACCGAAGCGAAAAAAGTGTCAACGATAAGATTTGGCCTGCGATC  
 421 GATAAAATCCTGCAGCAAAACGCCGAAACCTGTCCGATCCTGTCTGAGAAATGGTCTGGT  
 481 ATTACGTTAGCGGTGATCAGCTCAAGGGTCAGCGCCACAAACAGGAAGATCGCTTCTCTG  
 541 GCTTACCCTAACGGTCAATATATGGACCGTGGTGAGGACCCGATTTCTGTTCTGGCGGTT  
 601 TTCGATGGTCACGGCGGTACGAATGCTCTCAGTACGCGGCAGGTCTGTGGGAAACG  
 661 TGGCTGGAGGTGCGTAAAAGCCGCGACCCGTCTGACTCTCTGGAGGACCAGCTGCGCAAA  
 721 TCTCTCGAACTGCTGGACGAACGTATGACCGTGCGTTCTGTCAAAGAATGCTGGAAAGGC  
 781 GGTCTTACCGCTGTCTGCTGTGCAATCGACATGGACCAGAAGCTCATGGCGCTCGCATGG  
 841 CTGGGCGACTCTCCGGGCTATGTTATGTCTAACATCGAATTCCGTCACCTCACTCGTGGT  
 901 CATTCTCCTTCTGATGAGCGCGAAGCACGTCGCGTAGAGGAGGCAGGTGGCCAACTCTTT  
 961 GTAATCGGTGGTGAATCCGTGTTAACGGTGTGCTGAACCTGACGCGCGCACTCGGCGAT  
 1021 GTTCCGGGTGCTCCGATGATCTCTAACGAACCGGAAACTTGCCAGGTTCCGATCGAATCT

1081 AGCGACTATCTCGTGCTCCTGGCGTGCGATGGTATCAGCGACGTATTCAATGAGCGTGAC  
1141 CTGTATCAACTGGTCGAAGCGTTTGCCAATGATTACCCGGTTGAAGACTACGCGGAAGT  
1201 TCTCGCTTTATTTGCACCAAGGCCATTGAAGCAGGCTCTGCGGACAATGTATCTGTTGTT  
1261 ATTGGTTTCCTGCGTCCGCCGAGGACGTGTGGAACTGATGAAGCATGAATCTGACGAC  
1321 GAAGACTCTGATGTTACCGATGAGGAATAA

-----  
The oligonucleotide assembly is:  
-----

	1	10	20	30	40	50	60
1 --->							
1	ATGGA	AAAA	AGTTA	TGAAGA	ACGTG	ACGCGG	TGTTGAAGATCAC
							GCGCCACAAACTTCTAGTGTAGCCGCTGGCAGCG
	M	E	K	V	N	E	E
							R
	D	A	V	F	E	D	H
							I
							G
							D
							R
							R
3 --->							
61	gttcgtt	ctctg	ctcga	agaag	ctttc	gcggac	gaaatggagaaaac
	GCGAGACA	AGCAAG	AGACG	AGCTTCT			cctcttttggagaatg
	R	S	V	R	S	L	L
							E
							E
							A
							F
							A
							D
							E
							M
							E
							K
							T
							S
							Y
5 --->							
121	gacgtt						CCGCACATCCCAATCCGTTTTCGCCACCCGCCG
	ctgcaactcca	acgcctat	ggggtgtc	ggcggtg	taggggtaggc		
	D	V	E	V	A	D	T
							P
							Q
							P
							H
							I
							P
							I
							R
							F
							R
							H
							P
							P
7 --->							
181	ATCGCTGGCCCAGTTCATGATGTGTTT						cgacatctttcagaaaatg
	CGGGTCAAGTACTACACAAGCCACTACGATAGGTGCTGTAGAAAGTCTTTTAC						
	I	A	G	P	V	H	D
							V
							F
							G
							D
							A
							I
							H
							D
							I
							F
							Q
							K
							M
9 --->							
241	atgaaacgcggtcaggctgttgacttctgccactgggtgtc						
	TACTTTG						gaagacggtgaccacagagtggactaacgtgactt
	M	K	R	G	Q	A	V
							D
							F
							C
							H
							W
							V
							S
							H
							L
							I
							A
							T
							E
11 --->							
301	CGACGAAAAGTTTTCCGAAGTTGCGTTCCGTGACGTTTCAGTACAACCCGACATCTAC						
	tagctgcttttcaaaaaggcttca						TGTTGGGCCTGTAGATG
	I	D	E	K	F	S	E
							V
							A
							F
							R
							D
							V
							Q
							Y
							N
							P
							D
							I
							Y
13 --->							
361	GT						agcgaaaaaactgttcaacgataagatttggcctgcgatc
	CAATGCCTGAGATGATGGCTTTCGCTTTTTTGACAAGTTGCTAT						tag
	V	T	D	S	T	T	E
							A
							K
							K
							L
							F
							N
							D
							K
							I
							W
							P
							A
							I

```
|      |      |      |      |      |
13 --->
421 gataaaatcctgcagcaaaa      GATCCTGTCTGAGAAATGGTCTGGT
ctatTTtaggacgtcgtTTTtgcggctTTtgacaggctaggacagactctttaccaga
      <--- 12
      D K I L Q Q N A E T C P I L S E K W S G
|      |      |      |      |      |
15
481 ATTCACGTTAGCGGTGATCAGCTCAAGGGTCAGCG      tg
      AGTCGAGTTCACAGTCGCGGTGTTTGTCTTCTAGCGAAGGAC
      I H V S G D Q L K G Q R H K Q E D R F L
|      |      |      |      |      |
--->
541 gcttaccctaacgggtcaatatatggaccgtggtgaggacccgatttctgttctggcgg
CGAATGGGATTGCCAGT      ggctaaagacaagaccgccaa
      <--- 14
      A Y P N G Q Y M D R G E D P I S V L A V
|      |      |      |      |      |
17 --->
601 CGAATGCTCTCAGTACGCGGCAGGTCATCTGTGGGAAACG
aagctaccagtgccgccagtgcttacgagagtcatgcgc
      <--- 16
      F D G H G G H E C S Q Y A A G H L W E T
|      |      |      |      |      |
19 --->
661 TGGCTGGAGGTGCGTAAAAG      ggaccagctgcgcaaa
CCGACCTCCACGCATTTTCGGCGCTGGGCAGACTGAGAGACCTCCTGGTCGACGCGTTT
      <--- 1
      W L E V R K S R D P S D S L E D Q L R K
|      |      |      |      |      |
721 tctctcgaactgctggacgaacgtatgaccgtgcgttctgtcaa
A      ctggcagcgaagacagtttcttacgacctttccg
8
      S L E L L D E R M T V R S V K E C W K G
|      |      |      |      |      |
21 --->
781 CGCTGTCTGCTGTGCAATCGACATGGACCAGAAGCTCATGGCGCTCGCATGG
ccaagatggcgacagacgacacgtta      GCGTACC
      <--- 20
      G S T A V C C A I D M D Q K L M A L A W
|      |      |      |      |      |
23 --->
841 CTGGGCGA      catcgaattccgtcaactcactcgtggt
GACCCGCTGAGAGGCCCGATAACAATACAGATTGTAGCTTAAGGCAGTTGAGTG
      <--- 22
      L G D S P G Y V M S N I E F R Q L T R G
```

```

|      |      |      |      |      |      |
25 --
901 cattctccttctgatgagcgcgaagcacgtcg      TCTTT
      cgcgcttcgtgcagcgcatctcctccgtccaccggttgagaaa
      H S P S D E R E A R R V E E A G G Q L F
|      |      |      |      |      |
->
961 GTAATCGGTGGTGAACTCCGTGTTAACGGTGTGCTGAACCTGACGCGCGCACTCG
      cattagccaccacttga      ACTGCGCGCGTGAGCCGCTA
      <--- 24
      V I G G E L R V N G V L N L T R A L G D
|      |      |      |      |      |
27 --->
1021 tctctaacgaaccggaacttgccaggttccgatcgaatct
      CAAGGCCAGCAGGCTACTAGAGATTGCTTGGCCTTTGAA
      <--- 26
      V P G R P M I S N E P E T C Q V P I E S
|      |      |      |      |      |
29 --->
1081 agcgactatctctgtgctcc      GACGTATTCAATGAGCGTGAC
      tcgctgatagagcacgaggaccgcacgctaccatagtcgctgcataagttactcgactg
      <--- 28
      S D Y L V L L A C D G I S D V F N E R D
|      |      |      |      |      |
31
1141 CTGTATCAACTGGTCGAAGCGTTTGCCAATGATTACCCG      tg
      GCAAACGGTTACTAATGGGCCAACTTCTGATGCGCCTTGAC
      L Y Q L V E A F A N D Y P V E D Y A E L
|      |      |      |      |      |
--->
1201 tctcgctttatttgcaccaaggccattgaagcaggctctgcggaatgtatctgttg
      AGAGCGAAATAAACGTGGT      gacgcctgttacatagacaacaa
      <--- 30
      S R F I C T K A I E A G S A D N V S V V
|      |      |      |      |      |
33 --->
1261 CGCAGGACGTGTGGAACTGATGAAGCATGAATCTGACGAC
      taaccaaaggacgcaggcggcgctcctgcacaccttg      TG
      <--- 32
      I G F L R P P Q D V W K L M K H E S D D
|      |      |      |      |      |
1321 GAAGACTCTGATGTTACCG
      CTTCTGAGACTACAATGGCTACTCCTTATT
      <--- 34
      E D S D V T D E E X
-----
```

```

The total codon usage score ..... 0.000
The total length score ..... 0.000
The total melting temperature score ... 0.000
The total repeat score ..... 0.000
The total pattern score ..... 0.000
The total mispriming score ..... 0.000
The total AT content score ..... 0.000
The total GC content score ..... 0.000
The total fixed gap score ..... N/A
The OVERALL score ..... 0.000

```

## DETAILED CODON FREQUENCY REPORT

[ Codon, AA, Frequency, # of times used in coding sequence ]

```

-----
TTT F 0.29   7  TCT S 0.32  25  TAT Y 0.35   4  TGT C 0.39   2
TTC F 0.71  10  TCC S 0.27   1  TAC Y 0.65   7  TGC C 0.61   7
TTA L 0.03   0  TCA S 0.05   0  TAA X 0.63   1  TGA X 0.35   0
TTG L 0.06   0  TCG S 0.07   0  TAG X 0.08   0  TGG W 1.00   8

CTT L 0.06   0  CCT P 0.11   3  CAT H 0.30   4  CGT R 0.64  16
CTC L 0.08  11  CCC P 0.02   0  CAC H 0.70  10  CGC R 0.33  12
CTA L 0.01   0  CCA P 0.15   3  CAA Q 0.19   5  CGA R 0.01   0
CTG L 0.77  22  CCG P 0.72  15  CAG Q 0.81  13  CGG R 0.01   0

ATT I 0.34   7  ACT T 0.29   4  AAT N 0.17   4  AGT S 0.05   0
ATC I 0.66  17  ACC T 0.54   8  AAC N 0.83   8  AGC S 0.24   4
ATA I 0.01   0  ACA T 0.05   0  AAA K 0.79  14  AGA R 0.01   0
ATG M 1.00  11  ACG T 0.13   3  AAG K 0.22   6  AGG R 0.00   0

GTT V 0.40  20  GCT A 0.28   6  GAT D 0.46  15  GGT G 0.51  19
GTC V 0.14   3  GCC A 0.16   3  GAC D 0.54  28  GGC G 0.43   9
GTA V 0.20   4  GCA A 0.24   7  GAA E 0.75  31  GGA G 0.02   0
GTG V 0.27   8  GCG A 0.32  14  GAG E 0.25  11  GGG G 0.04   0

```

Frequency Range	Number of Codons
0% - 4%	0
5% - 9%	11
10% - 14%	9
15% - 19%	15
20% - 24%	32
25% - 29%	30
30% - 34%	58
35% - 39%	26
40% - 44%	9
45% - 49%	15
>= 50%	245

Total Codons Used = 450

Tm Range	# of Overlaps
<60	0
60	0
61	16
62	17
63	0
64	0
65	0
66	0
67	0
>=68	0

-----  
Tm Range = 2.0

Overlap Len Range # of Oligos

-----  
 <16 3  
 16 0  
 17 2  
 18 5  
 19 7  
 20 3  
 21 7  
 22 3  
 23 2  
 24 0  
 25 0  
 >=26 1  
 -----

Lowest Overlap = 15

Length Range # of Oligos

-----  
 < 50 2  
 50- 51 0  
 52- 53 0  
 54- 55 0  
 56- 57 0  
 58- 59 0  
 60- 61 32  
 62- 63 0  
 64- 65 0  
 66- 67 0  
 68- 69 0  
 >= 70 0  
 -----

Longest = 60

Sequence Patterns Screened (As Supplied By User)

-----  
 None found  
 -----

34 oligonucleotides need to be synthesized

-----  
 1 ATGGAAAAAGTTAATGAAGAACGTGACGCGGTGTTTGAAGATCAC 45  
 2 TCTTCGAGCAGAGAACGAACAGAGCGGCGACGGTCGCCGATGTGATCTTCAAACACCGCG 60  
 3 GTTCGTTCTCTGCTCGAAGAAGCTTTCGCGGACGAAATGGAGAAAACCTCTTACGACGTT 60  
 4 CGGATTGGGATGTGCGGCTGTGGGGTATCCGCAACCTCAACGTCGTAAGAGGTTTTCTCC 60  
 5 CCGCACATCCCAATCCGTTTTTCGCCACCCGCCGATCGCTGGCCAGTTCATGATGTGTT 60  
 6 GTTTCATCATTTTTCTGAAAAGATGTCGTGGATAGCATACCCGAACACATCATGAACTGGGC 60  
 7 CGACATCTTTCAGAAAAATGATGAAACGCGGTGAGGCTGTTGACTTCTGCCACTGGGTGTC 60  
 8 ACTTCGGAAGAACTTTTCGTCGATTTTCAGTCGCAATCAGGTGAGACACCCAGTGGCAGAAG 60  
 9 CGACGAAAAGTTTTCCGAAGTTGCGTTCCGTGACGTTTCAGTACAACCCGGACATCTACGT 60  
 10 TATCGTTGAACAGTTTTTTCGCTTCGGTAGTAGAGTCCGTAACGTAGATGTCCGGGTTGT 60  
 11 AGCGAAAAAACTGTTCAACGATAAGATTTGGCCTGCGATCGATAAAATCCTGCAGCAAAA 60  
 12 AGACCATTTCTCAGACAGGATCGGACAGGTTTCGGCGTTTTGCTGCAGGATTTTATCGAT 60  
 13 GATCCTGTCTGAGAAATGGTCTGGTATTCACGTTAGCGGTGATCAGCTCAAGGGTCAGCG 60  
 14 TGACCGTTAGGGTAAGCCAGGAAGCGATCTTCTGTTTGTGGCGCTGACCCTTGAGCTGA 60  
 15 TGGCTTACCTAACGGTCAATATATGGACCGTGGTGAGGACCCGATTCTGTTCTGGCGG 60  
 16 CGCGTACTGAGAGCATTCTGTGACCGCCGTGACCATCGAAAACCGCCAGAACAGAAATCGG 60  
 17 CGAATGCTCTCAGTACGCGGCAGGTCATCTGTGGGAAACGTGGCTGGAGGTGCGTAAAG 60  
 18 ATTTGCGCAGCTGGTCTCCAGAGAGTCAGACGGGTCGCGGCTTTTACGCACCTCCAGCC 60  
 19 GGACCACTGCGCAAATCTCTCGAACTGCTGGACGAACGTATGACCGTGCGTTCTGTCAA 60  
 20 ATTGCACAGCAGACAGCGGTAGAACCCTTTCCAGCATTCTTTGACAGAACGCACGGTC 60  
 -----

60  
60  
60  
60  
60  
60  
60  
60  
60  
60  
60  
60  
60

## 1 SOLUTION

#	Tm	Len	Score	TmRange	Short	Long	#Olig	#Repeat	#Misprime
1	62	60	0.000	2.0	15	60	34	0	0

Helix Systems -- Center for Information Technology  
http://helix.nih.gov  
National Institutes of Health, Department of Health and Human Services  
DNAWorks Web Site: http://helixweb.nih.gov/dnaworks